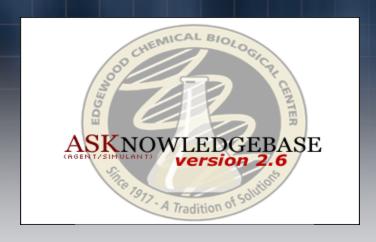
US Army Edgewood Chemical Biological Center Chemical Biological Agent Simulant Knowledgebase (ASK) (Information Workbench & Graphical User Interface)



Raymond E. Jablonski and William P. Ashman
US Army ECBC, RDECOM
David Mark Harrah, Matthew Williamson
Oak Ridge Institute for Science & Education (ORISE)

Edgewood Chemical Biological Center 5183 Blackhawk Road, ATTN: AMSRD-ECB-RT-IM Aberdeen Proving Ground, Maryland, USA 21010-5424



Email: ray.jablonski@us.army.mil

Phone: (410) 436-3566 FAX: (410) 436-2742

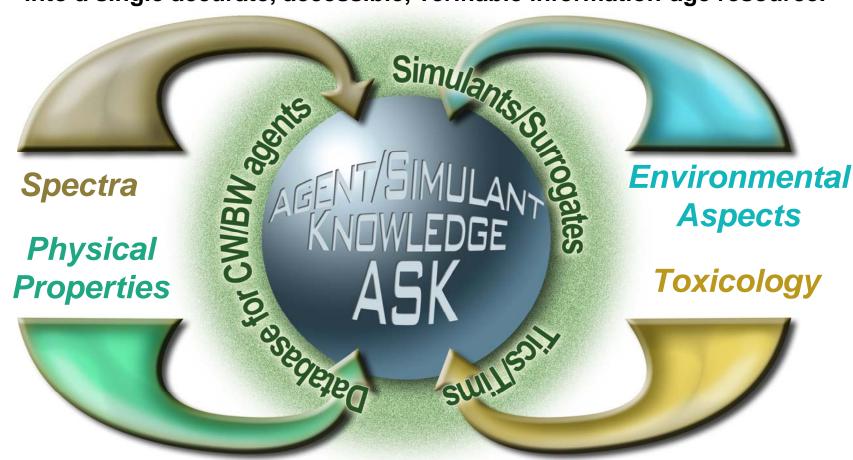
maintaining the data needed, and c including suggestions for reducing	lection of information is estimated to ompleting and reviewing the collect this burden, to Washington Headqu uld be aware that notwithstanding ar OMB control number.	ion of information. Send comments a arters Services, Directorate for Infor	regarding this burden estimate of mation Operations and Reports	or any other aspect of th , 1215 Jefferson Davis l	is collection of information, Highway, Suite 1204, Arlington	
1. REPORT DATE 17 NOV 2004		2. REPORT TYPE N/A		3. DATES COVE	RED	
4. TITLE AND SUBTITLE				5a. CONTRACT NUMBER		
US Army Edgewood Agent Simulant Ki	0	5b. GRANT NUMBER				
Graphical User Interface)				5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)				5d. PROJECT NUMBER		
				5e. TASK NUMBER		
				5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Edgewood Chemical Biological Center 5183 Blackhawk Road, ATTN: AMSRD-ECB-RT-IM Aberdeen Proving Ground, Maryland, USA 21010-5424				8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)		
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAIL Approved for publ	LABILITY STATEMENT ic release, distributi	on unlimited				
	OTES 49, 2004 Scientific C Cland on 15-17 Nove					
14. ABSTRACT						
15. SUBJECT TERMS						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON	
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified	UU	42	ALSI ONSIBLE I ERSON	

Report Documentation Page

Form Approved OMB No. 0704-0188

- Overview
- ASK Module Summary
 - Chemical, Biological, Simulant Application, Environmental, Spectral
- Current ASK GUI capabilities
 - Browse
 - Search / Match
 - Properties Calculator
 - Save/Print Options
 - Data Viewers
- Lessons Learned / Current Obstacles
- Future Plans

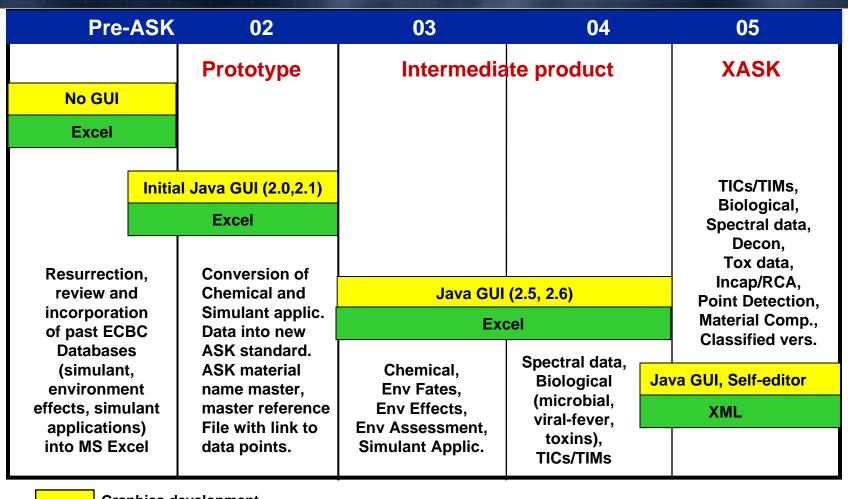
GOAL: Collect, verify, and integrate available agent/simulant knowledge into a single accurate, accessible, verifiable information-age resource.



ASK Development Approach

- <u>Assess</u> the current status of chemical and biological agent/ simulant information, query CB community for CB data needs/requirements
- <u>Collate</u> available agent/simulant data (identify, obtain, review, verify, incorporate information into database); identify key data gaps
- Develop an accurate, verifiable, readily available central <u>data</u>
 <u>repository and software "workbench" tool</u> for CB data mining
 - ☐ Verified/Standardized data and other CB data
 - ☐ Capability: Search, Analyze, and Report
 - ☐ Design based on user and data requester feedback
- Establish a central joint service CB <u>simulant advisory service</u> to provide CB Agent/Simulant information and expertise
- <u>Provide data and technical expertise</u> to the Joint Services and servicespecific counterparts (e.g., Standard Simulants for Chemical Agent Detector Calibration Study, US Army Natick Protection Against Toxic Industrial Chemicals (PATIC) initiative)
- <u>Provide information and technical expertise</u> to numerous defense and civilian agencies involved in chemical and biological defense activities (e.g., DHS, DOE NARAC, EPA, NIOSH)

ASK Development History



Graphics development

Data Collection

Accomplishments to Date

- O ASK v2.6 Current Version
- O JAVA Based Software Utility/ Data Mining Tool and User Interface.
- O Chemical, Simulant Applications, Biological (Microbial, Viral-Fever, Toxins), Spectral Data and Environmental Fate/Effects modules currently under development and in use by ECBC.
- O Due to user feedback and data requests, other modules are being considered. Some initial development has been conducted for modules to address: Decontamination, Material Compatibility, Point Detection, Smoke-Obscurants-Interferents, Demilitarization, Diseases Caused by Materials in the Environment, Toxicological data, Incapacitation/Riot Control Agents, Source Term.
- O Original data source references are being acquired, and as the original data source becomes available it is reviewed by subject matter experts.
- O In FY04 ASK has responded to ~75 requests for data/information.

- Overview
- ASK Module Summary
 - Chemical, Biological, Simulant Application, Environmental, Spectral
- Current ASK GUI capabilities
 - Browse
 - Search / Match
 - Properties Calculator
 - Save/Print Options
 - Data Viewers
- Lessons Learned / Current Obstacles
- Future Plans

Chemical Module

Purpose: Data source for physical, chemical, thermodynamic, HazMat, and toxicological information for chemical agents, simulants (surrogates), precursors, degradation and by-products, Toxic Industrial Chemicals (TICs)/Toxic Industrial Materials (TIMs), Interferents, Tracers/Taggants and other misc. chemicals

Data Collected: ~1600 materials, ~180 Data Fields

Source of Data: Published reports, chemistry reference books, ECBC lab notebooks, journal articles, other database efforts, some web sources (e.g., NIST/TOXNET)

Status: Ongoing effort. FY05 focus on TICs/TIMs, precursors and degradation/by-products information, possibly also include some pesticides.

Name: Biological - Microbial Module

Purpose: Data source of bacteria material. This module with include both "classic" bio warfare agents, bacteria commonly found in the environment and bacteria and/or bio-aerosol simulants.

Data Collected: Over 500 microbial materials (agents, simulants and various strains) identified, 172 data fields

Source of Data: Published reports, reference books, journal articles, past database and literature search efforts, some web sources.

Status: Ongoing. Initiated in late FY03 through first half of FY04. Additional data/information remains to be incorporated into ASK in FY05.

Name: Biological - Viral-Fever Module

Purpose: Data source of viruses which cause disease in humans. This module with include both "classic" bio warfare viruses, viruses commonly found in the environment and viral simulants.

Type of Data Collected: ~50 viruses/fevers identified. ~150 data fields

Source of Data: Published reports, reference books (Virus Taxonomy: The Classification and Nomenclature of Viruses), journal articles, other database efforts, some web sources (International Committee on Taxonomy of Viruses database).

Status: Initiated late in FY04. To be completed in FY05.

Name: Biological - Toxins Module

Purpose: Data source of toxins of biological origin and their simulants. This module with include both "classic" bio warfare toxins, toxins/venoms found in the environment and toxin simulants. Toxin categories include bacterial toxins, fungal toxins, plant toxins, dinoflagellate toxins, coelenterata toxins, mollusca toxins, insecta toxins, amphibia toxins, and reptilia venoms.

Data Collected: ~180 toxins identified, 100 data fields

Source of Data: Published reports, reference books, journal articles, past database and literature search efforts, some web sources.

Status: Initiated late in FY04. To be completed in FY05.

Name: Simulant Applications Module

Purpose: Identify information that describes and/or evaluates agents/simulants used for research, acquisition and test & evaluation applications. Main goal is to identify the agents & simulants used, and identify references. If available, identify what agent was simulated, for what application, simulant strong points and/or deficiencies. Provide specific information on document reference and evaluations from simulant users, if possible.

Data Collected: Over 2000 agent/simulant applications identified for ~500 materials. Four main module categories (Identifier, Application, Reference, Simulant User Info) with a total of **29 data fields**.

Source of Data: Published reports, journal articles, simulant user survey.

Status: Ongoing effort. In FY05 plan to expand on current module to also include agent test information.

Name: Environmental Module

Purpose: Provide information to address factors that effects CB materials in the environment, including environmental effects and environmental fate. Also gather and incorporate information required for developing test environment assessments (e.g., environmental impact statement, environmental assessment, testing permits).

Data Collected: ~200 Materials (Simulants, Agents, Interferents, Tracers and Taggants), 100 Data Fields

Source of Data: DoD T&E environmental assessment/impact documents, published reports, journal articles, environmental fate assessments.

Status: First phase of consolidating past environmental databases, data sets completed. Continue to work with ECBC, DPG, WSMR and other T&E locations. Once published, plan to incorporate into ASK the DPG report "Environmental Assessment For Test Materials to be used in Laboratory, Chamber, and Field Tests at US Army Dugway Proving Ground, Dugway, Utah."

Name: Spectral Data Viewer & Module

Purpose: Create a consolidated set of verified spectral data for CB defense acquisition, research, test and evaluation use. Reduce duplicate data sets. Identify and eliminate invalid data sets.

Data Collected: Infrared, Mass Spec, Raman, Electromagnetic, UV/Visible, Nuclear Magnetic Resonance (NMR), Fluorescence, Others. Have collected over **1600 spectral data files**.

Source of Data: Have incorporated data from ECBC, NIST, Pacific Northwest National Lab (PNNL) and Dugway Proving Ground (DPG).

Status: Current focus on data verification. Simple data viewer that allows user to preview data, since user will use more sophisticated software to analyze spectral data. Text output to be available to the user in x-y coordinates with experimental conditions identified. Still at a point of initial development, not releasing data. Pointing requesters to originator of data.

Other ASK modules under consideration

- Decontamination
- Material compatibility
- Incapacitation Agent
- Point Detection
- Smoke, Obscurants & Interferents
- Demilitarization
- Diseases Caused by Materials in the Environment
- Toxicological data
- Source term

- Overview
- ASK Module Summary
 - Chemical, Biological, Simulant Application, Environmental, Spectral
- Current ASK GUI capabilities
 - Browse
 - Search / Match
 - Properties Calculator
 - Save/Print Options
 - Data Viewers
- Lessons Learned / Current Obstacles
- Future Plans

Browse capability:

Allows the user to quickly select a material and view it's data.

Search capability:

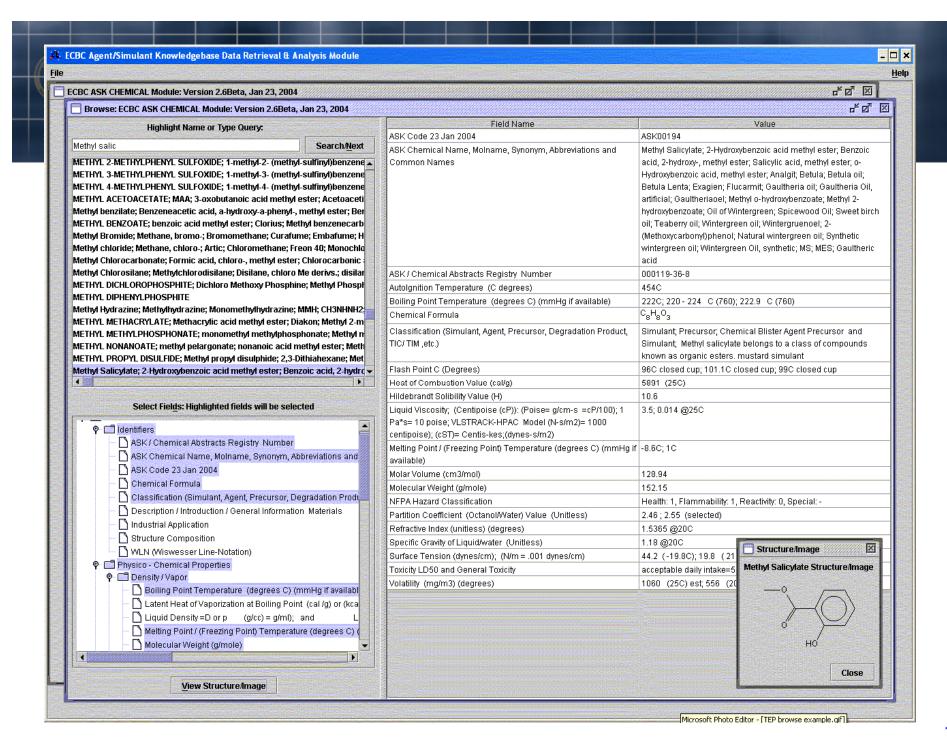
Allows the user to search by data field, whether by identifier or data value(s). Can query by specific data value or data ranges for individual or multiple data fields. Searching by data value(s) results in a 'relevance' ranking of the material(s) shown in the search results.

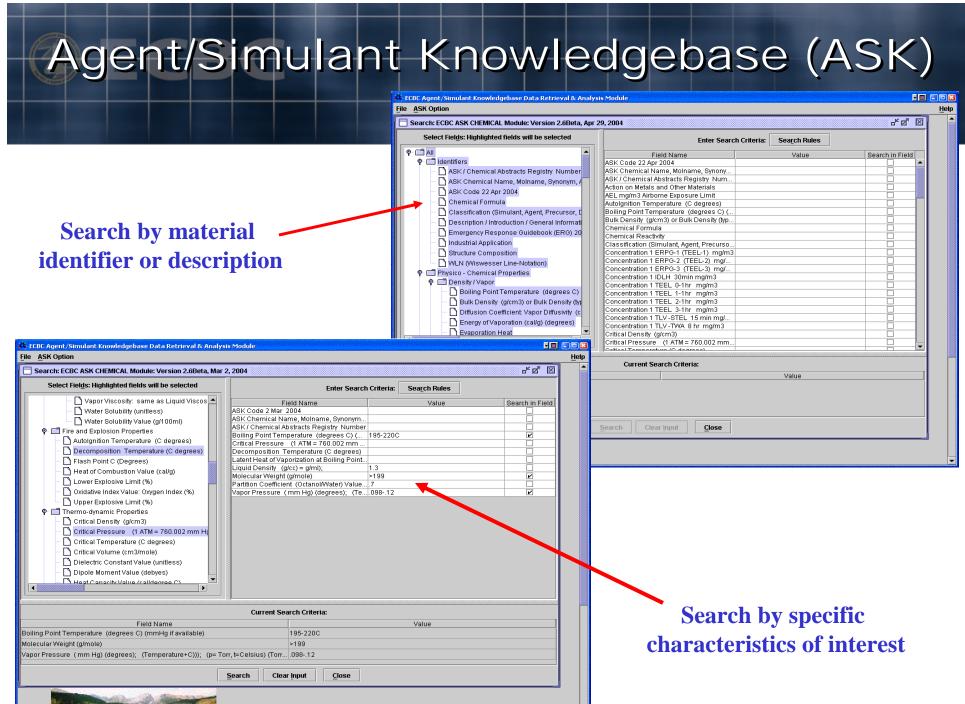
Match capability:

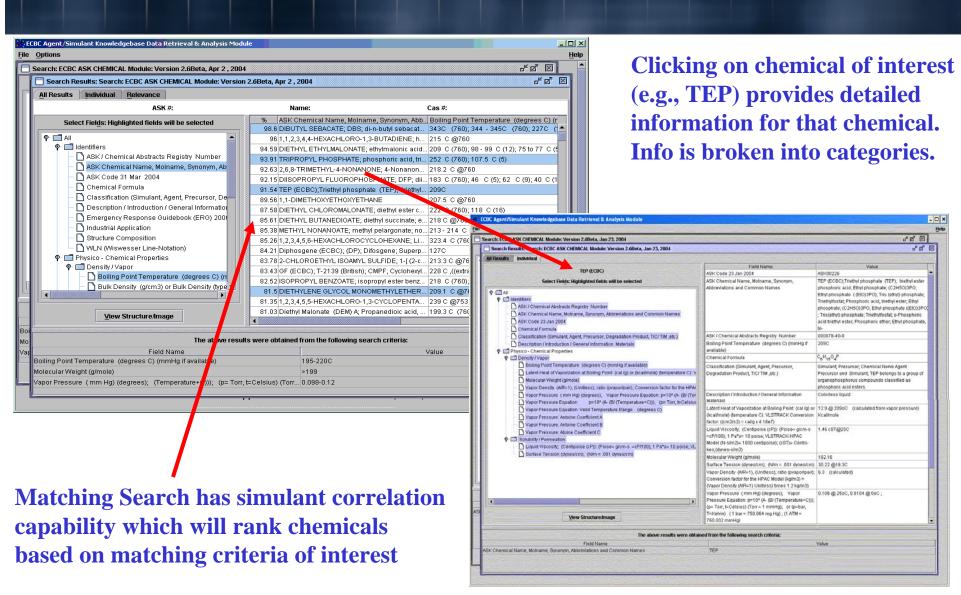
This option uses ASK's internal data values for searching. Allows the user to directly select an ASK material, select specific data field(s) and perform an ASK search to identify materials based on user's criterion. As in the Search capability, the user is provided a 'relevance' ranking of the materials shown in the search results.

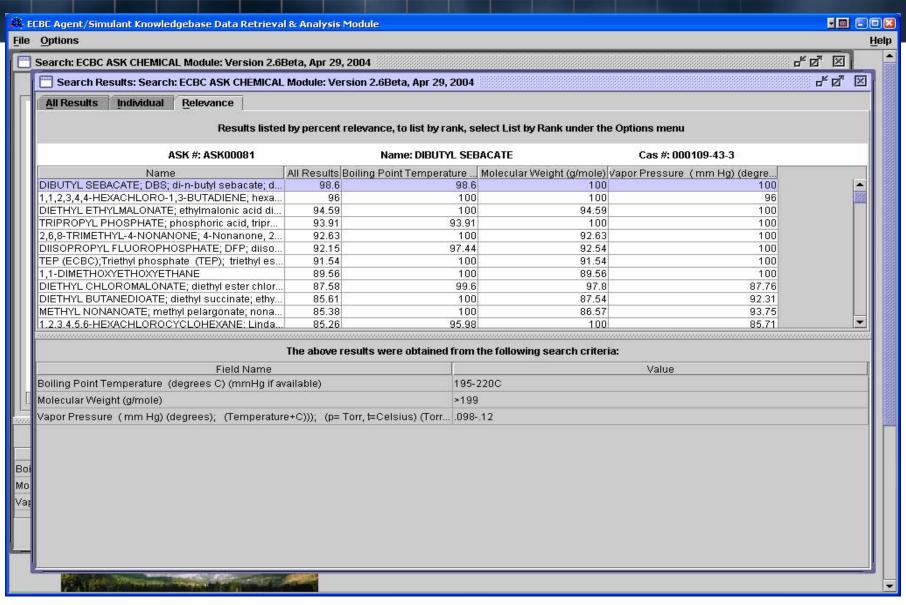
Property Calculator:

Calculates property value(s) using equations incorporated into ASK. Allows the user to directly calculate physical properties at various variable experimental conditions (i.e., temperature).

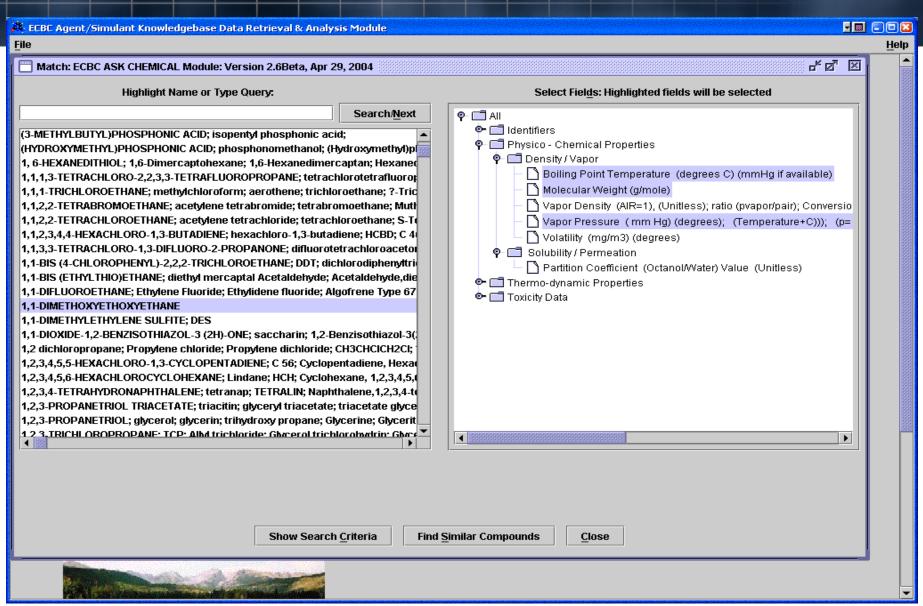




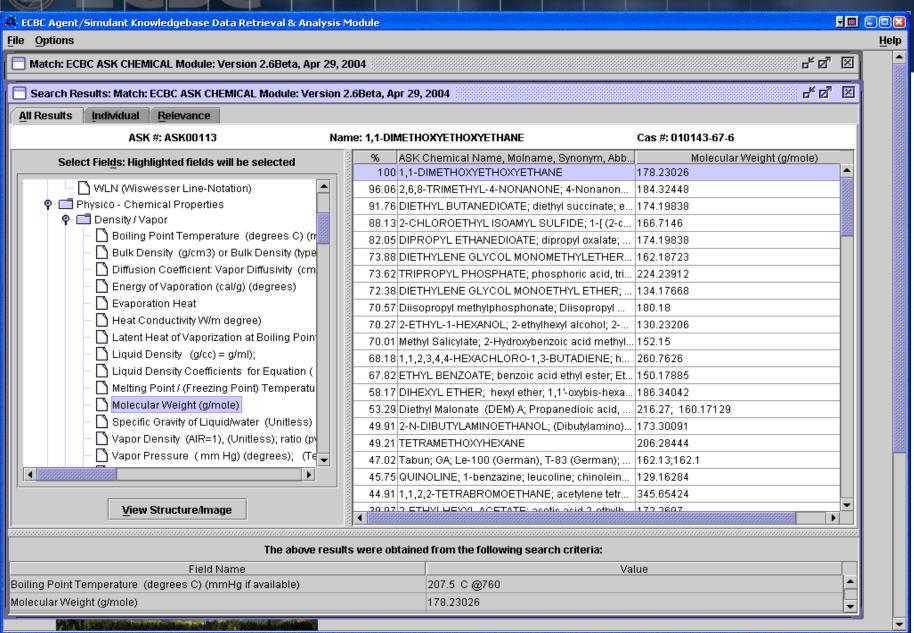




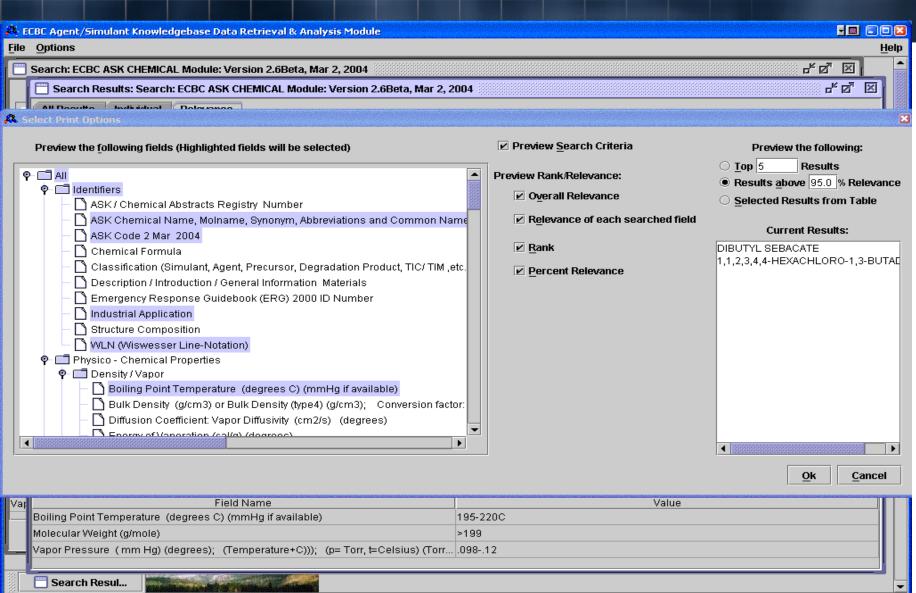
DECEASK Match Option



ASK Match Result



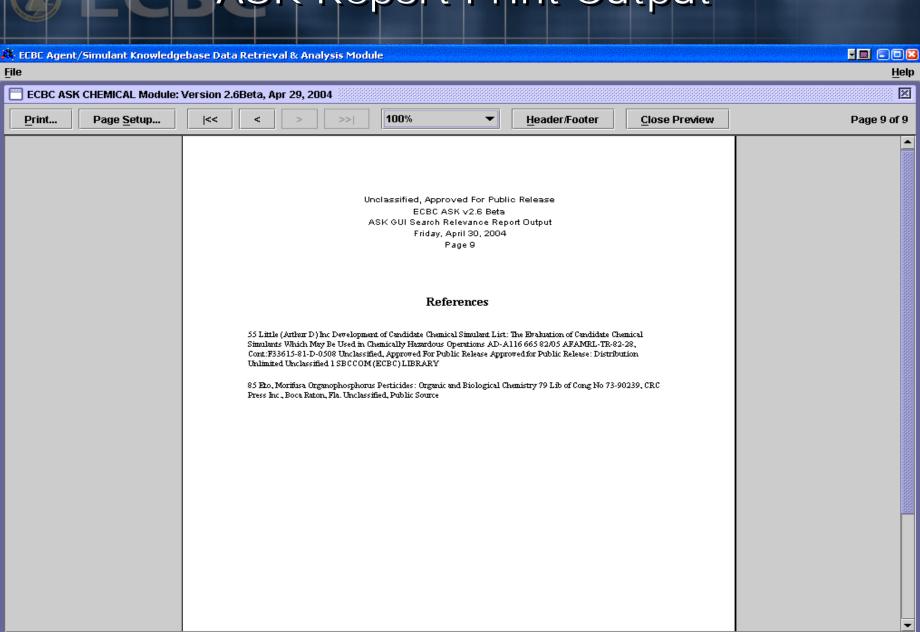
FASK Report Print Output

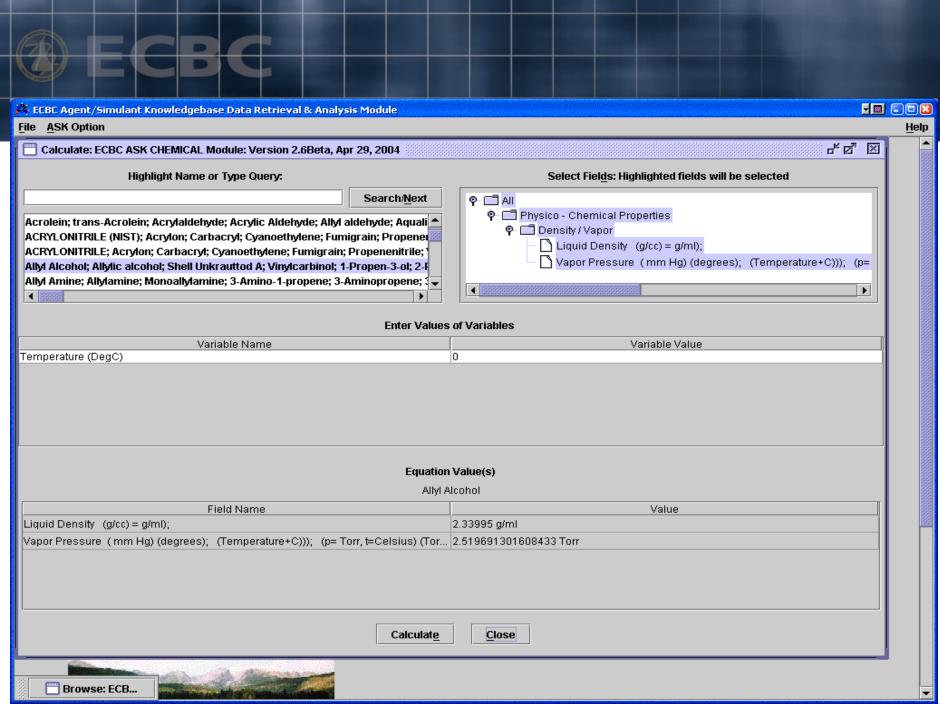


ASK Print Header /Footer Option

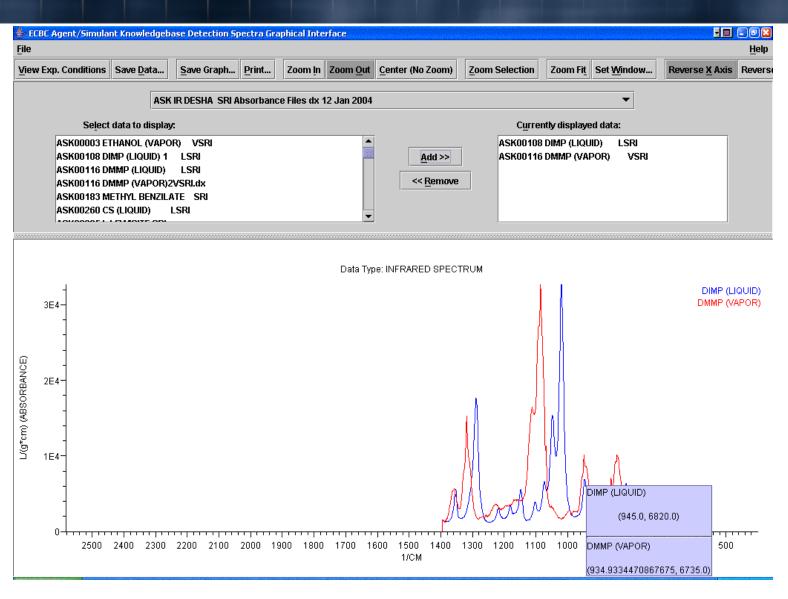
🙉 ECBC Agent/Simulant Knowledg	gebase Data Retrieval & Analysis Module						
<u>F</u> ile		<u>H</u> elp					
ECBC ASK CHEMICAL Module: Version 2.6Beta, Apr 29, 2004							
Print Page <u>S</u> etup	<	Page 1 of 9					
		A 1970					
	Header/Footer	A Header/Footer					
		1. Select what to add to the header or footer:					
	Unclassified, Approved For Public Release ECBC ASK v2.6 Beta Page Number Date	○ <u>T</u> ext					
	ASK GUI Search Relevance Report Output Friday, April 30, 2004						
	Page 1						
	Search: ECBC ASK CHEMICAL Module: Version 2.6Be April 30, 2004						
	○ Friday, April 30	, 2004					
	Search Criteria						
		2. To add: Drag to a slot in the header or footer To remove: Drag from slot to surrounding area Page ###					
	Boiling Point Temperature (degrees C)(mmHg if available)						
	Molecular Weight (g/mole)	Header					
	Vapor Pressure (mm. Hg) (degrees); (Temperature+C))); (p= Torr,t=Celsius) (Torr = 1 mm.Hg T=Kelwin) (1 bar = 750.064 mg Hg); (1 ATM = 760.002 mm.Hg) ASK GUI Search Relevan	ce Report Output					
		Friday, April 30, 2004					
	Page ##	#					
	1,1-DIMETHOXYETHOXYETHANE Footer						
	н,с ^ о						
	"" 0 ~ 0 ~ 0 ~ 0 ~ 0 ~ 0 ~ 0 ~ 0 ~ 0 ~ 0						
	Clear	<u>C</u> los					
	Overall Rank: 8 SECRC Agent/Simulant Wuntitled - Paint						
	Overall Percent Relevance: 89.56						
	73.045.						
	Identifiers						

ECBASK Report Print Output





Agent/Simulant Knowledgebase (ASK) Spectral Data Module



Cother ASK Features

- Mouse key identification of reference with data point
- References summarized with report printout or HMTL
- ASK report header/footer can be edited by user
- User can specify want data to save/print
- Embedded images, reports, tables option
- Individual column search option
- Relevance scores can be reviewed individually
- GUI internal help/error messages
- ASK workbench search strategy

- Overview
- ASK Module Summary
 - Chemical, Biological, Simulant Application, Environmental, Spectral
- Current ASK GUI capabilities
 - Browse
 - Search / Match
 - Properties Calculator
 - Save/Print Options
 - Data Viewers
- Lessons Learned / Current Obstacles
- Future Plans

ASK - Lessons Learned / Current Obstacles

- Should "any" data be included in database
 - □ Past Database Efforts
 - □ OGA Database Efforts
- Should database have started from scratch instead of inheriting past databases as a starting point? Past databases included good data, but also mistakes and (in some cases) no references.
- Better data definition standard at start of database construction
- Better data evaluation audit trial
- Stricter data quality control/rules
- Obtaining permission to use international and intelligence data for classified version of ASK
- Including copyrighted information
- Portion of ASK data is Unclassified, but DoD only

- Overview
- ASK Module Summary
 - Chemical, Biological, Simulant Application, Environmental, Spectral
- Current ASK GUI capabilities
 - Browse
 - Search / Match
 - Properties Calculator
 - Save/Print Options
 - Data Viewers
- Lessons Learned / Current Obstacles
- Future Plans

FCBASK Project FY05

ASK v3.0 (9/05) Database: Continue data additions, verification and Platform Upgrades

- Based on CB User discussions and queries, continue to populate current data fields for the <u>primary</u> set of modules (Chemical, Biological, Spectral, Environmental, Simulant Applications)
- Continue work on secondary set of modules as time/funding permits
- Complete classified version of ASK
- Acquire, validate and incorporate CB data to include ECBC, Dugway, literature, experimental and other government agencies (OGA's) and international CB agent/ simulant data
- Upgrade ASK on-line user interface and Graphical User Interface Platform

ASK Data Structure Problems with the exported Excel flat file process

- No built-in validity checking of the data
- Adding or modifying functionality to the application can require extensive, cumbersome changes to the data file
- The file must be manually edited in Excel and then exported in a different format
- While the "Excel" approach was simple & fast, there
 is little flexibility, extensibility or power in using this
 method for data collecting and mining

ASK Data Structure Why the need to convert to XML

- Improves modularity, flexibility, consistency, efficiency and ease of use of product
 - Development
 - Maintaining the software
 - Extending the functionality of product
 - End User
- Advantages of XML vs. Relational Model (Development & Maintenance)
 - XML is better for semi-structured data
 - XML is independent of any proprietary software (Excel, Oracle)
 - Proprietary software has associated costs, learning curve for both the developer and user
 - Proprietary software adds a step of complexity to whole process
 - XML is simply text (can outlive the application; readable, structured format that other applications can use)
 - A relational database requires a proprietary, binary format (dependent on that software)
- Currently, entire data from module is read into memory.
 - Simple & fast, but little flexibility, extensibility or power in the approach

ASK Data Structure Why XML? Advantages in the application

- Data Entry
 - Allows for direct editing of data, no separate data editor required
 - Can provide extensive validity checking of data entered
 - ☐ Greatly improves database consistency
- Development
 - Interchangeable tools already exist (Open Source)
 - ☐ Read, Write (Relaxer)
 - □ Data Query (eXist)
- Maintenance/Extensibility
 - The meaning of each piece of data is fully described
 - Modifications/extensions to application have little reason to modify the data/format
- End User
 - Application is more powerful and intuitive
 - Searching possibly exporting results as PDF
 - More flexibility with references

ASK "Success Stories"

- Completed version 2.6 of ASK more of a "work bench" software package
- Completed spectra data GUI module. GUI incorporates CB user recommended data formatting, data graphic display, comparison and data format requirements
- Conducted CB user surveys on bio and environmental modules data field requirements. Designed/incorporated user requirements into ASK modules.
- Identified, collected, incorporated and verified data for all ASK modules.
- Leveraged with various agencies to share data and established data collaboration efforts among the data information gathering community.

ASK data/information contributors

Chemical module

ECBC

DPG

DOT&E (Applied Resources, Inc.)

NSWC DTRA

CBIAC-Battelle

WMSR

Wright-Patterson AF Base

SAIC NIST EPA

Army Corps of Engineers

Academia

International data sources

Biological module

ECBC DPG DOT&E NSWC DTRA

CBIAC-Battelle

MRIID DTIC-BPL Academia

International data sources

Spectra Data Module

ECBC DPG PNNL NIST CBIAC Academia

Simulant Applications

ECBC DPG DOT&E

Wright-Patterson AF Base

DTIC

Environmental module

ECBC DPG DOT&E WMSR DTRA NIST EPA

Army Corps of Engineers

ASK Advisory Project

- ASK is planned to be available to US Department of Defense,
 U.S. Government Agencies and their qualified Contractors
- Contact Information:
 - Mr. William P. Ashman (410) 436-3430
 - william.ashman@sbccom.apgea.army.mil
 - Mr. Raymond Jablonski (410) 436-3566
 - ray.jablonski@sbccom.apgea.army.mil
 - Address: U.S. Army ECBC

ATTN: AMSSB-RRT-IM

5183 Blackhawk Road

APG, MD 21010-5424



Edgewood Chemical Biological Center 5183 Blackhawk Road, ATTN: AMSRD-ECB-RT-IM

Aberdeen Proving Ground, Maryland, USA 21010-5424



Email: ray.jablonski@us.army.mil

Phone: (410) 436-3566 FAX: (410) 436-2742

ASK Data Base Future Efforts

Current Obstacles in Spectra data collection/evaluation

- Different experimental conditions
 - Need to compare data sets and determine which are actually the same (Not as easy as it sounds)
- Why?
 - Experimental conditions not recorded
 - Consistency in scaling (units)
 - Transmittance vs. absorbance
- Data collected directly from lab equipment producing invisible characters, thus those causing problem with software reading data file

ASK "Success Stories"

- ASK Project's efficiency. Project has already saved a number of organizations/project time and money in their search for data.
 ASK has streamlined the process of data searching, initial data analysis, agent-to-simulant comparison and providing the information in a user friendly summary format.
 - DTRA/SAIC personnel came to ECBC and used ASK for obtaining WSMR Preliminary Environmental Impact Statement CB environmental data (5 days)
 - Intelligent Automation, Inc (AF) request for spectra for use in CB spectra identification algorithm development (~1 hr)
 - Plasmasol Corp (CRADA) need for CB physicochemical and simulant recommendations for decon studies (3 hrs)
 - Numerous requests and use of ASK by CBIAC to quickly respond to CB information requests (~20min each)
- Have averaged 60 to 80 requests a year from approximately 30 different agencies/organizations
- DoD Modeling (e.g., JEM), DHS/DOE modeling and NATO have expressed interested in using ASK as a standardized data resource